

Durham Research Online

Deposited in DRO:

10 March 2017

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Wolff, N. and Kemter, K. and Schweinberger, S.R. and Wiese, H. (2014) 'What drives social in-group biases in face recognition memory? ERP evidence from the own-gender bias.', *Social cognitive and affective neuroscience.*, 9 (5). pp. 580-590.

Further information on publisher's website:

<https://doi.org/10.1093/scan/nst024>

Publisher's copyright statement:

This is a pre-copyedited, author-produced version of an article accepted for publication in *Social Cognitive and Affective Neuroscience* following peer review. The version of record Nicole Wolff, Kathleen Kemter, Stefan R. Schweinberger, Holger Wiese; What drives social in-group biases in face recognition memory? ERP evidence from the own-gender bias. *Social Cognitive and Affective Neuroscience* 2014; 9 (5): 580-590 is available online at: <https://doi.org/10.1093/scan/nst024>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

What drives social in-group biases in face recognition memory?

ERP evidence from the own-gender bias

Nicole Wolff ^{1,2}, Kathleen Kemter ¹, Stefan R. Schweinberger ^{1,2} and Holger Wiese ^{1,2}

¹ Friedrich Schiller University of Jena, Germany

² DFG Research Unit Person Perception

Corresponding author:

Nicole Wolff

Person Perception Research Unit

Friedrich Schiller University of Jena

Leutragraben 1

07743 Jena

Germany

E-Mail: Nicole.Wolff@uni-jena.de

Phone: +49 3641 945940

Fax: +49 3641 945182

Abstract

It is well established that memory is more accurate for own- relative to other-race faces (own-race bias, ORB), which has been suggested to result from larger perceptual expertise for own-race faces. Previous studies also demonstrated better memory for own- relative to other-gender faces which is less likely to result from differences in perceptual expertise, and rather may be related to social in- versus out-group categorization. We examined neural correlates of the own-gender bias using event-related potentials (ERP). In a recognition memory experiment, both female and male participants remembered faces of their respective own gender more accurately as compared to other-gender faces. ERPs during learning yielded significant differences between the subsequent memory effects (subsequently remembered – subsequently forgotten) for own- as compared to other-gender faces in the occipito-temporal P2 and the central N200, whereas neither later subsequent memory effects nor ERP old/new effects at test reflected a neural correlate of the own-gender bias. We conclude that the own-gender bias is mainly related to study phase processes, which is in line with socio-cognitive accounts.

Words: 172

Keywords: own-gender bias, ERP, face processing, Dm effect, old/new effect

Introduction

Although humans can recognize a massive number of previously encountered faces, this ability varies with the category a specific face belongs to. For example, participants are more accurate at remembering faces of their own relative to other ethnic (for a review, see Meissner & Brigham, 2001) or age groups (see e.g. Bartlett & Fulton, 1991; Bartlett & Leslie, 1986; for a recent meta-analysis see Rhodes & Anastasi, 2012). Both these effects, commonly referred to as *own-race* and *own-age* biases respectively, have been explained by differences in contact and/or perceptual expertise with own-group and other-group persons. For instance, most Caucasian participants in European countries spend more time with people from their own ethnic group than with people from another ethnic group. Similarly, most young adults have substantially more contact to their own age group relative to elderly persons. Accordingly, it was suggested that the face representational system (the so-called face space), being shaped by experience, is normally tuned to differentiate optimally between individual faces of a viewer's own age and ethnic group, but less well for other-group faces (Furl, Phillips, & O'Toole, 2002; Valentine, 1991, 1992; Byatt & Rhodes, 2004). Such differences in the representation of own- and other-group faces have been suggested to underlie biases in face recognition memory.

Similar to these well-established own-race and own-age biases, an *own-gender* bias in face recognition memory has also been described. However, previous results have not been entirely consistent. While women are commonly found to be more accurate at remembering female as compared with male faces (Lewin & Herlitz, 2002; Lovén, Herlitz, & Rehnman, 2011), results for male participants appear more variable, either showing (i) a corresponding own-gender bias with better memory for male than female faces (Wright & Sladden, 2003), (ii) similar performance for male and female faces (Lewin & Herlitz, 2002), or (iii) even enhanced memory performance for female relative to male faces (McKelvie, Standing, St

Jean, & Law, 1993). Although an own-gender bias in women was suggested to result from enhanced perceptual expertise (Lewin & Herlitz, 2002; Lovén et al., 2011; Rehnman & Herlitz, 2006), empirical support for this suggestion is lacking, as no measure of contact or expertise was reported in those studies. By contrast, and irrespective of the above-described discrepancies, it could be argued that any own-gender bias in face memory is unlikely to result from different perceptual expertise or daily-life contact, to the extent that most people in western societies regularly interact with persons from both genders.

Alternatively, biases in face recognition memory have been explained by socio-cognitive accounts, emphasizing the importance of categorizing a face as belonging to either a social “in-group” or “out-group” (Hugenberg, Young, Bernstein, & Sacco, 2010; Sporer, 2001). Accordingly, face memory biases may occur not because of differences in expertise or contact, but because faces perceived as belonging to a social “out-group” are processed at a categorical level, whereas “in-group” faces are individualized. Such individualization is assumed to result in more in-depth processing, and thus more accurate recognition memory. Supporting these ideas, it was demonstrated that even “in-group” faces with respect to race and age, when arbitrarily labelled as belonging to the participants’ university, were remembered more accurately as similar faces when labelled as being affiliated to a different university (Bernstein, Young, & Hugenberg, 2007). Importantly, socio-cognitive accounts assume that face memory biases reflect differential processing of in- versus out-group faces during learning. Relevant empirical support comes from observations that both the own-race bias and biases based on labelling same-race faces as belonging to a social out-group disappeared when participants were briefed about memory biases, and were instructed to pay close attention to out-group faces during learning (Hugenberg, Miller, & Claypool, 2007; Young, Bernstein, & Hugenberg, 2010).

In sum, while an own-gender bias is not easily integrated into contact- or expertise-based accounts, it can be explained by socio-cognitive accounts, given that other-gender faces

are perceived as belonging to a social “out-group”. It also remains unclear whether similar or different processes underlie the own-race, own-age, and own-gender biases. Thus, even when the own-gender bias may not be related to expertise, it remains possible that the own-race bias is. It is therefore important to identify and compare cognitive and neural processes underlying these effects. Event-related potentials (ERP) provide an excellent method to examine this issue.

While behavioural measures can only capture the end-product of a sequence of mental sub-processes initiated after stimulus presentation, ERPs provide detailed chronometric information about those sub-processes. The first ERP component usually examined in face perception experiments is the P1, which is maximal over occipital sites at approximately 80-130 ms after stimulus onset. P1 is elicited by any visual stimulus, and its amplitude and latency are known to be highly sensitive to physical characteristics, including spatial frequency, luminance and contrast (Luck, 2005).

More importantly, *face* stimuli elicit a large negative deflection maximal at right occipito-temporal channels after approximately 170 ms. This so-called N170 (Bentin et al., 1996; Bötzel et al., 1995; Eimer, 2011) was suggested to reflect the structural encoding of faces (Bentin et al., 1996) or the detection of a face-like stimulus in the visual field (Schweinberger & Burton, 2003). While several studies did not find differences for own- versus other-race faces in N170 (Caldara et al., 2003; Wiese, Stahl, & Schweinberger, 2009), a number of recent studies reported larger N170 for other- as compared to own-race faces (Caharel et al., 2011; Stahl, Wiese, & Schweinberger, 2008, 2010; Walker, Silvert, Hewstone, & Nobre, 2008; Wiese, 2012). This discrepancy may be partly explained by task demands: Senholzi and Ito (in press) demonstrated larger N170 for other-race faces only when participants were asked to individuate the faces. A number of the above-cited studies used recognition memory paradigms, in which participants were to memorize faces, which likely induced a motivation to process these on an individual level. Other studies examined effects

of face ethnicity on the central P200 (also typically peaking around 170 ms), which likely reflects the positive counterpart of the N170, the vertex positive potential (Jeffreys, 1989), and observed larger VPP amplitudes for other-race faces (Ito & Bartholow, 2009; Ito & Urland, 2003). Moreover, N170 has been found to be larger for old as compared to young faces (Wiese, Komes, & Schweinberger, 2012; Wiese, Schweinberger, & Hansen, 2008). By contrast, previous studies on processing facial gender did not find differences in N170 amplitude for male versus female faces (Mouchetant-Rostaing & Giard, 2003; Mouchetant-Rostaing, Giard, Bentin, Aguera, & Pernier, 2000). Interestingly, however, female participants have been reported to demonstrate smaller N170 amplitudes in gender categorization (female/male) as compared to orientation tasks (face to the left/right), while male participants yielded no comparable pattern (Sun, Gao, & Han, 2010). Together, although preliminary evidence suggests some (task-dependent) differences between male and female observers, there is no evidence for overall differences in N170 depending on facial gender.

The occipito-temporal P2 represents a positive deflection around 200-280 ms. P2 was found to be decreased for vertically stretched versus normal faces (Halit, De Haan, & Johnson, 2000) and was suggested to reflect the processing of metric distances between facial features (Latinus & Taylor, 2006; Mercure, Dick, & Johnson, 2008). Larger P2 responses have been observed for own- versus other-race faces (Stahl et al., 2008, 2010), for young versus old faces (Wiese et al., 2008), and for veridical versus spatially caricatured faces (Kaufmann & Schweinberger, 2012). Accordingly, P2 may reflect the perceived typicality of a face.

The subsequent negative component, the N250, is consistently larger for repeated relative to non-repeated faces in immediate repetition priming experiments (the so-called N250r effect (Schweinberger, Pfütze, & Sommer, 1995; Schweinberger, Pickering, Jentsch, Burton, & Kaufmann, 2002)). More recently, N250 has also been associated with face learning, since more negative N250 amplitudes for learned vs. novel faces have been reported

(Kaufmann, Schweinberger, & Burton, 2009; Tanaka, Curran, Porterfield, & Collins, 2006). The N250 may reflect the access to perceptual representations of individual faces, and thus the processing of facial identity. In addition, generally more negative amplitudes in the N250 time range were reported for other- as compared to own-race faces (Stahl et al., 2010; Wiese, 2012), and effects of stimulus repetition and face category have been shown to add up in the N250 (Wiese, Kaufmann, & Schweinberger, in press). Accordingly, effects of face repetition and face category are independent and may represent different processes. More negative amplitudes in this time range have also been observed for emotional faces (Schupp et al., 2007), which has been interpreted as “tagging” of particularly salient faces for further processing. Similar to effects in the N170, face category effects in N250 may be related to task demands: Ebner and colleagues (2011) found more negative N250 for own-age faces in a gender categorization task, while studies using recognition memory paradigms found more negative N250 for other-age faces (Wiese, 2012). To the best of our knowledge, no previous study systematically examined effects of face gender on P2 or N250.

In addition, several ERP effects related to episodic memory have been identified. For instance, in the learning phases of memory experiments, words later remembered at test elicit more positive amplitudes than later forgotten words (e.g., Paller, McCarthy, & Wood, 1988; Paller, Kutas, & Mayes, 1987). This so-called *Difference due to subsequent memory* (or Dm) effect is typically broadly distributed and sustained roughly between 300 and 1000 ms (e.g. Herzmann & Curran, 2011). Importantly, Dm effects were also described for faces (Sommer, Heinz, Leuthold, Matt, & Schweinberger, 1995; Sommer, Schweinberger, & Matt, 1991), and were recently observed to be larger for own- as compared to other-race faces (Lucas, Chiao, & Paller, 2011). In that study, a differential Dm effect for own- and other-race faces was detected in the occipito-temporal P2 (and polarity-reversed in the central N200), with more positive amplitudes for subsequently forgotten relative to subsequently remembered own-race faces, and with the opposite pattern for other-race faces. Herzmann and co-workers

(Herzmann, Willenbockel, Tanaka, & Curran, 2011) examined Dm effects for own- and other-race faces, which were either subsequently recollected (i.e., elicited recollection of study phase detail at test (see Yonelinas, 2002), or subsequently familiar (i.e., recognized at test without accompanying study phase detail). At some variance with the above-discussed findings, subsequently recollected other-race faces elicited more positive amplitudes than subsequently familiar other-race faces in Caucasian participants. In addition, no significant differences between subsequently recollected versus subsequently familiar own-race faces were detected. This was interpreted as representing similarly deep encoding for subsequently recollected and subsequently familiar own-race faces, but more elaborate learning of subsequently recollected other-race faces.

Finally, retrieval effects from episodic memory are typically detected in so-called ERP old/new effects during test phases of recognition memory experiments (see e.g., Rugg & Coles, 1995). Old/new effects consist of more positivity for correctly remembered items (hits) as compared to correctly rejected new items (correct rejections). They typically occur between 400 and 800 ms, with a left parietal maximum for verbal material (see e.g., Rugg & Curran, 2007). Previous studies on own-group biases reported larger centro-parietal old/new effects for own- relative to other-race faces (Herzmann et al., 2011; Stahl et al., 2010; Wiese, 2012), and for own- versus other-age faces in young participants (Wiese et al., 2012, 2008). Moreover, Guillem and Mogg (2005) found the old/new effect for faces to be reduced in male as compared to female participants. However, to the best of our knowledge no previous study directly investigated neural correlates of the own-gender bias.

We aimed at filling this gap by examining ERP correlates of recognition memory for female and male faces in both female and male participants. Although the present study did not test different biases in a combined experiment, a comparison of such ERP effects with previously observed neural correlates of the own-race bias may help to clarify the extent to which these biases rely on similar processes. While combining memory biases within the

same experiment facilitates direct comparison between biases (see e.g., Wiese, 2012), biases likely interact and are therefore not measured in their “purest” form in combined experiments. Accordingly, testing these biases in separate experiments represents important additional information.

We reasoned that both the absence of an effect that is consistently observed for other biases, and the observation of a novel ERP correlate specific to the own-gender bias would point to partly different processes underlying the respective memory phenomena. We considered the following specific hypotheses with respect to the ERP components under investigation: (i) Since other-race faces elicit a larger N170 in recognition memory studies (Stahl et al., 2010), and since this effect is correlated with the own-race bias in memory (Wiese, Kaufmann, & Schweinberger, in press), a similar finding for other-gender faces would suggest that face category effects in N170 are not based on perceptual expertise - as the processing of own- and other-gender faces likely does not differ with respect to expertise-based processing. (ii) P2 amplitude is larger for own- versus other-race faces, but participants with enhanced contact towards other-race faces do not show this effect (Stahl et al., 2008; Wiese et al., in press). Accordingly, assuming that face category effects in P2 are based on perceptual expertise, no differential P2 was predicted for own- versus other-gender faces. (iii) N250 has been observed to be more negative for other- as compared to own-race faces, and during test this difference was shown to correlate with the own-race bias in memory (Wiese et al., in press). If a corresponding effect were observed for the own-gender bias, it would likely not reflect perceptual expertise-based processing. Finally, with respect to memory-related ERP effects the following predictions were tested: Dm effects during learning were found to be larger, or to differ in scalp distribution, for own- and other-race faces (Lucas et al., 2011). Moreover, old/new effects at test were found to be larger for own- compared to other-race faces (Herzmann et al., 2011; Stahl et al., 2010). The presence of similar effects in the own-gender bias would point to an interpretation in terms of social categorization.

Methods

Participants

Participants were 28 female (18-33 years, mean age = 22.89 +/- 3.65 *SD*) and 28 male (18-32 years, mean age = 23.82 +/- 3.36 *SD*) undergraduate students from the University of Jena. All were right-handed according to a modified version of the Edinburgh Handedness Inventory (Oldfield, 1971), reported normal or corrected-to-normal vision, and received a honorarium of 5.00 €/h or course credits. None reported neurological or psychiatric disorders, or to receive central-acting medication.

Stimuli

Stimuli were 120 unfamiliar female faces aged 18-40 years ($M = 24.09 \pm 3.16$ *SD*) and 120 male faces aged 19-40 years ($M = 25.70 \pm 4.74$ *SD*). All images depicted colour portraits with frontal views and neutral expressions. Stimuli were taken from the CAL/PAL (Minear & Park, 2004), the FACES (Ebner, Riediger, & Lindenberger, 2010), the FERET (Phillips, Wechsler, Huang, & Rauss, 1998), and the GUFID databases (Burton, White, & McNeill, 2010). Using Adobe PhotoshopTM, faces were cut out (i.e., no clothing or background visible) and pasted to a uniform black background. Stimuli were cropped to a frame of 299×380 pixels (10.55 × 13.41cm), resulting in a visual angle of 6.7° × 8.5° at a viewing distance of 90 cm, and presented on a computer monitor using E-PrimeTM.

Procedure

Main experiment. Participants were seated in a dimly lit, electrically shielded and noise-attenuated chamber (400-A-CT-Special, Industrial Acoustics, Niederkrüchten, Germany) with their heads in a chin rest. The experiment consisted of a practice block (16

trials) and 4 experimental blocks. Each block was divided into a study and a test phase. In individual trials, a fixation cross was first displayed for 500 ms, followed by a face for 5000 ms (study phases) or for 2000 ms (test phases). Each trial ended with a blank screen (500 ms).

During each study phase, 30 faces (50% female) were presented, and participants were instructed to memorize and categorize the faces by gender. Study and test phases were separated by a 20 s break. During the ensuing test phase, all 30 faces from the directly preceding study phase were presented in randomized order together with 30 new faces (again 50% female). Participants decided via button presses with their left or right index finger whether a face had been presented in the directly preceding study phase (“learned”) or not (“new”). Speed and accuracy were emphasized. If participants classified a face as “learned”, a remember/know rating with an additional “guess” option (Gardiner & Richardson-Klavehn, 2000) was conducted. Between blocks, participants were allowed a self-paced period of rest. Key assignment and allocation of stimuli to learned or non-learned conditions were balanced across participants.

Responses during test phases were sorted into four conditions, for female and male faces separately: hits (correctly identified learned faces), misses (learned faces incorrectly classified as new), correct rejections (CR, new faces correctly identified as new), and false alarms (FA, new faces incorrectly classified as learned). Measures of sensitivity ($d' = z[\text{hits}] - z[\text{FA}]$) and response bias ($C = -0.5 [z(\text{hits}) + z(\text{FA})]$) were calculated according to signal detection theory (see e.g., Green & Swets, 1966). Additionally, a memory bias score, consisting of the difference between d' for female and male faces divided by the sum of the two measures, was calculated. Finally, percentages of ‘remember’ and ‘know’ responses for hits relative to the number of learned items were analyzed for male and female faces.

Face Rating. After the experiment, each participant rated those 120 faces presented during learning for distinctiveness (from 1 = not distinctive to 6 = very distinctive) and

attractiveness (from 1 = not attractive to 6 = very attractive). Faces were presented in randomized order, and participants were instructed to rate them spontaneously.

Contact Questionnaire. Finally, all participants completed a questionnaire, which asked for the amount of contact (in hours/week), number of contact persons (per week) and contact quality (from 1 = very superficial to 4 = very intense) for female and male people in daily-life situations (such as job/university, meeting friends/spare time activities, family, and domestic circumstances (see Wiese, 2012). Total scores were calculated for each participant by summing up (hours/week, number of persons/week) or averaging (contact quality) self-report measures from the different situations separately for contact towards male and female persons.

EEG Recording and analysis

EEG was recorded using a 32-channel BioSemi Active II system (BioSemi, Amsterdam, Netherlands). Active sintered Ag/AgCl-electrodes were mounted in an elastic cap with recording sites at Fz, Cz, Pz, Iz, FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, F9, F10, FT9, FT10, TP9, TP10, P9, P10, PO9, PO10, I1 and I2. EEG was recorded continuously (512 Hz sampling rate, DC to 155 Hz). Note that BioSemi systems work with a “zero-Ref” setup with ground and reference electrodes replaced by a so-called CMS/DRL circuit (cf. to <http://www.biosemi.com/faq/cms&drl.htm> for further information).

Blink artefacts were corrected using the algorithm implemented in BESA 5.1 (Berg & Scherg, 1994). EEG was segmented from –200 until 1000 ms relative to face onset, with 200 ms pre-stimulus baseline. Trials contaminated by non-ocular artefacts and saccades were rejected from further analysis using the BESA 5.1 tool, with an amplitude threshold of 100 μ V and a gradient criterion of 75 μ V. Artefact-free trials were recalculated to average reference, averaged according to experimental condition and digitally low-pass filtered at 40 Hz (12 db/oct, zero phase shift). Learning phase data were averaged separately for

subsequently remembered and subsequently forgotten male and female faces. Data from the test phases were averaged separately for hits and CR of male and female faces. An inclusion criterion of at least 16 trials per condition per participant was applied. For learning phase data, three participants (two female and one male, all exhibiting fewer than 16 trials in one of the ‘subsequently forgotten’ conditions) did not reach this criterion. In order to keep the design counterbalanced with respect to key and stimulus assignments to experimental conditions when excluding these participants, two additional female and three additional male participants were randomly selected and excluded. Thus, 24 male and 24 female participants entered the statistical analyses of ERP data from the learning phases. Analysis of test phase data was based on all 28 participants per group. The maximum number of trials to be achieved, during learning and test, per participant and condition was 60. More specifically, during learning, the mean numbers of attained trials contributing to each averaged ERP of female participants was for subsequently remembered female faces: $M = 39.71 \pm 8.09 SD$, for subsequently remembered male faces: $M = 34.83 \pm 10.42 SD$ for subsequently forgotten female faces: $M = 20.25 \pm 5.75 SD$ and for subsequently forgotten male faces: $M = 23.79 \pm 8.45 SD$. For male participants the corresponding numbers were for subsequently remembered female faces: $M = 35.13 \pm 8.34 SD$, for subsequently remembered male faces: $M = 36.63 \pm 7.88 SD$, for subsequently forgotten female faces: $M = 23.50 \pm 7.32 SD$, and for subsequently forgotten male faces: $M = 21.83 \pm 7.43 SD$. During test mean numbers of attained trials for female participants were for hits to female faces: $M = 40.11 \pm 6.21 SD$, for CR to female faces: $M = 45.43 \pm 9.21 SD$, for hits to male faces: $M = 36.32 \pm 8.56 SD$ and for CR to male faces: $M = 46.43 \pm 9.16 SD$. For male participants the corresponding numbers were for hits to female faces: $M = 36.50 \pm 7.12 SD$, for CR to female faces: $M = 48.46 \pm 6.43 SD$, for hits to male faces: $M = 38.71 \pm 6.38 SD$ and for CR to male faces: $M = 48.11 \pm 7.51 SD$.

Early ERP components (P1, N170) were measured at the electrodes of their respective maximum relative to a 200 ms prestimulus baseline. Mean P1 amplitudes were analyzed at O1 and O2 between 105 and 145 ms. Mean amplitudes of N170 and two following components (P2, N250) were analyzed at P9 and P10 in consecutive time windows (N170: 155 -195 ms; P2: 220 – 270 ms; N250: 270 – 400 ms). Finally, mean amplitudes for Dm effects in N200 (220 -270 ms), the late Dm (450- 1000 ms) and old/new effects (400 – 675 ms) were calculated at F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4. Separate mixed-model ANOVAs were performed for each component during both learning and test. When appropriate, degrees of freedom were corrected according to the Huynh-Feldt procedure.

Results

Performance

Accuracies during learning were near ceiling (see Table 1) and not statistically tested. To analyze reaction time (RT) data from the learning phases (see Table 1), a mixed-model ANOVA with the within-subject factors ‘face gender’ (male faces, female faces) and the between-subjects factor ‘participant gender’ (male participants, female participants) was conducted, which yielded no significant results (all $p > .05$).

- Please enter Table 1 about here –

Memory performance during test phases was analyzed with a corresponding mixed-model ANOVA on d' (see Table 1, Figure 1), which yielded a significant interaction of ‘participant gender x face gender’ ($F[1,54] = 12.29, p < .01$). Separate analyses for each participant group revealed superior performance for own- as compared to other-gender faces for both female ($F[1,27] = 7.12, p < .05$) and male participants ($F[1,27] = 5.19, p < .05$).

- Please enter Figure 1 about here -

A corresponding ANOVA on C revealed a significant interaction of ‘face gender x participant gender’ ($F[1,54] = 8.91, p < .01$). Post hoc tests demonstrated female participants to respond more conservatively for male as compared to female faces ($F[1,27] = 9.53, p < .01$). Male participants showed no significant effect ($F[1,27] = 1.06, p > .05$).

During test, analysis of remember-responses for hits revealed an interaction of ‘face gender x participant gender’ ($F[1,54] = 4.13, p < .05$), reflecting more remember-responses for male relative to female faces in male participants ($F[1,27] = 7.53, p < .05$), but no difference in female participants ($F < 1$). A mixed-model ANOVA on know-responses for hits yielded no significant effects (all $p > .05$).

Face Ratings

Attractiveness. An analogous mixed-model ANOVA on attractiveness ratings yielded main effects of ‘face gender’ ($F[1,238] = 4.23, p < .05$), indicating higher attractiveness ratings for female vs. male faces (see Table 2), and ‘participant gender’ ($F[1,238] = 74.58, p < .001$), reflecting overall higher ratings by female vs. male participants. Furthermore, a significant interaction of ‘face gender x participant gender’ ($F[1,238] = 13.68, p < .001$) was observed. Post hoc tests revealed a significant main effect of ‘face gender’ for female participants ($F[1,119] = 10.02, p < .01$), with higher ratings for female vs. male faces, but no corresponding effect for male participants ($p > .05$).

Distinctiveness. A corresponding ANOVA on distinctiveness ratings yielded main effects of ‘face gender’ ($F[1,238] = 3.90, p < .05$), indicating higher ratings for male vs. female faces (see table 2), and ‘participant gender’ ($F[1,238] = 12.67, p < .001$), reflecting overall higher ratings by male vs. female participants. A significant interaction of ‘face gender x participant gender’ ($F[1,238] = 15.30, p < .001$) was also observed. Post hoc tests

revealed no significant effect of ‘face gender’ for female participants ($p > .05$), whereas male participants exhibited significantly higher distinctiveness ratings for male vs. female faces ($F[1,119] = 13.82, p < .001$).

Effects of attractiveness and distinctiveness on memory. To test for potential effects of the above described differences between male and female participants on face memory, we conducted an ANCOVA on d' with the within-subjects factor ‘face gender’ and the between-subjects factor ‘participant gender’, with the covariates ‘ Δ attractiveness’ (female – male faces) and ‘ Δ distinctiveness’ (female – male faces). Replicating the results described above, this analysis yielded an interaction of ‘participant gender x face gender’ ($F[1,52] = 9.94, p < .01$), indicating higher d' for own- as compared to other-gender faces for both participant groups. Neither of the two covariates yielded a significant effect (both $p > .05$). Furthermore, in additional analyses neither ‘ Δ attractiveness’ nor ‘ Δ distinctiveness’ correlated significantly with the memory bias score (both $p > .05$). Accordingly, differences between groups in perceived distinctiveness or attractiveness of male and female faces can not explain the own-gender bias.

Contact Measures (see Table 3)

A mixed-model ANOVA on *contact time* (in hours/week) with the within-subjects factor ‘contact persons’ gender’ and the between-subjects factor ‘participant gender’ yielded no significant effects (all $p > .05$). A corresponding ANOVA on *number of contact persons* yielded a main effect of ‘participant gender’ ($F[1,54] = 12.52, p < .001$), indicating that male participants reported higher numbers of contact persons than female participants. A mixed-model ANOVA on *contact quality* revealed a significant interaction of ‘contact persons’ gender x participant gender’ ($F[1,54] = 5.48, p < .05$), reflecting more intense contact towards female vs. male contact persons in female participants ($F[1,27] = 6.79, p < .05$), but no significant effect of ‘contact persons’ gender’ in male participants ($p > .05$).

Importantly, the memory bias score did not correlate with contact as measured in hours per week, number of contact persons, or quality of contact (all $r < .1$, all $p > .05$).

- Please enter Table 3 about here -

Event-related potentials during learning (see Figure 2)

Please note that ERP effects reported here focus on those results directly relevant for the aims of the present study. Analysis of P1 and additional results for all remaining ERP components can be found in the supplementary material.

Unless stated otherwise, learning phase ERPs were analyzed by calculating mixed-model ANOVAs with the within-subject factors ‘face gender’, ‘subsequent memory’ (subsequently remembered, subsequently forgotten), and ‘hemisphere’ (left, right), as well as the between-subjects factor ‘participant gender’. An ANOVA on N170 amplitudes at P9/P10 revealed a significant main effect of ‘participant gender’ ($F[1,46] = 7.37, p < .01$), with larger N170 for female vs. male participants, as well as a significant interaction of ‘hemisphere x face gender x participant gender’ ($F[1,46] = 4.29, p < .05$). Post-hoc analyses for male participants yielded a significant effect of ‘face gender’ at electrode P9 ($F[1,23] = 5.48, p < .05$), reflecting larger N170 amplitudes for female vs. male faces, with no significant effect at P10 ($F < 1$). Female participants did not exhibit significant effects of ‘face gender’ in the N170 (both $p > .05$). The difference in N170 amplitude between female and male faces correlated neither with the differences between female and male faces in attractiveness or distinctiveness ratings, nor with the differences in contact measures towards female versus male persons (all $p > .05$).

A mixed-model ANOVA on P2 amplitudes revealed a significant three-way interaction of ‘subsequent memory x face gender x participant gender’ ($F[1,46] = 8.20, p < .01$). Post-hoc tests for female participants revealed significantly *less* positive amplitudes for

subsequently remembered male as compared to subsequently forgotten male faces ($t[23] = -2.16, p < .05$), while no significant difference for subsequently remembered versus forgotten female faces was detected ($p > .05$). For male participants, the post-hoc analysis revealed significantly *more* positive amplitudes for subsequently remembered versus forgotten male faces ($t[23] = 2.14, p < .05$). Again, no significant difference was obtained for subsequently remembered versus forgotten female faces ($p > .05$). In addition, the P2 amplitude difference between female and male faces was found to be significantly correlated with the difference in rated distinctiveness between female and male faces ($r = -.31, p < .05$) at electrode P10, reflecting more positive P2 amplitudes for female faces when female faces were rated as less distinctive, and more positive P2 amplitudes for male faces when male faces were rated as less distinctive. None of the remaining correlations were significant (all $p > .05$).

An ANOVA on N200 at left, midline, and right frontal, central and parietal electrodes, with repeated-measures factors ‘site’ (frontal, central, parietal) and ‘laterality’ (left, midline, right; replacing the ‘hemisphere’ factor), revealed an interaction of ‘subsequent memory x face gender x participant gender’ ($F[1,46] = 7.79, p < .01$). Post hoc tests in female participants revealed less negative amplitudes for subsequently remembered vs. subsequently forgotten male faces ($t[23] = 2.64, p < .05$), but no significant difference for female faces ($p > .05$). In male participants, subsequently remembered male faces were more negative than subsequently forgotten male faces ($t[23] = -2.18, p < .05$), while again no significant effect was observed for female faces ($p > .05$).

Analysis of N250 amplitudes at P9/P10 revealed no significant effects of interest (all $p > .05$). Similar to N170, the difference in N250 between female and male faces correlated neither with the difference between female and male faces in attractiveness or distinctiveness ratings, nor with the difference in contact measures towards female versus male persons (all $p > .05$).

Finally, between 450 -1000 ms, a mixed-model ANOVA at left, midline, and right frontal, central and parietal electrodes revealed a clear Dm effect of ‘subsequent memory’ ($F[1,46] = 9.57, p < .01$), with more positive amplitudes for subsequently remembered as compared to subsequently forgotten faces. Importantly, there were no significant with ‘participant gender’ or ‘face gender’ (all $p > .05$).

- Please enter Figure 2 about here -

Event-related potentials during test phases (see Figure 3)

Unless stated otherwise, test phase ERPs were analyzed using mixed-model ANOVAs with the within-subject factors ‘face gender’, ‘response type’ (hits, CR), and ‘hemisphere’ (left, right), as well as the between-subjects factor ‘participant gender’.

Analysis of N170 amplitudes yielded no significant effects (all $p > .05$). A mixed-model ANOVA on P2 amplitudes revealed a main effect of ‘face gender’ ($F[1,54] = 5.20, p < .05$), with more positive amplitudes for female as compared to male faces. There was an interaction of ‘hemisphere x response type x participant gender’ ($F[1,54] = 4.17, p < .05$), although separate post-hoc tests at P9 and P10 for male and female participants revealed no significant effects of ‘response type’ (all $p > .05$).

A corresponding ANOVA on N250 amplitudes resulted only in a main effect of ‘response type’ ($F[1,54] = 5.43, p < .05$), with more negative amplitudes for hits than CR (all other effects $p > .05$). Correlation analyses revealed no significant effects for the difference in ERP amplitudes between female and male faces in N170, P2, and N250 and the corresponding difference in attractiveness, distinctiveness, or contact ratings.

An old/new effect, with more positive amplitudes for hits than correct rejections, was observed between 400-675 ms (see Figure 3). A mixed-model ANOVA with repeated measurements on ‘site’, ‘laterality’, ‘face gender’, and ‘response type’, as well as the

between-subjects factor ‘participant group’, revealed a main effect of ‘response type’ ($F[1,54] = 62.40, p < .001$), with more positivity for hits versus CRs. Importantly, a three-way interaction of ‘response type x face gender x participant gender’ ($F[1,54] = 4.04, p < .05$) was additionally observed. A post hoc test indicated larger old/new effects for female as compared to male faces ($F[1,27] = 5.12, p < .05$) in female participants. Although a similar own-gender effect appears to be present in the waveforms of male participants at central electrodes (Figure 3), a post hoc test did not indicate significant differences in old/new effects for female and male faces in male participants ($p > .05$).

Finally, we calculated an analysis which aimed at examining recollection-based recognition memory in isolation. A corresponding mixed model ANOVA with “remember”-responses versus CR as levels of the ‘response type’ factor revealed a trend for a significant main effect of ‘response type’ ($F[1,54] = 3.24, p = .077$).

- Please enter Figure 3 about here -

Discussion

The present study is the first that examined neural correlates of an own-gender bias in face memory. We observed more accurate recognition performance for own- relative to other-gender faces, in both female and male participants. Analysis of ERP components related to perceptual face processing did not reveal strong evidence for different processing of own- and other-gender faces. More specifically, neither N170, nor P2 and N250 yielded systematic differences between own- and other-gender faces. However, the analysis of memory-related effects in the occipito-temporal P2 and the central N200 during learning revealed differential effects of subsequent memory for male faces in female and male participants, and therefore a potential neural correlate of the own-gender bias. Finally, old/new effects at test were larger

for own- as compared to other-gender faces in the female participant group. These findings are discussed below against previous results from the own-race and own-age bias.

The role of expertise versus social category membership for the OGB

As detailed in the introduction, the majority of previous studies found an OGB in female, but not male participants, who additionally demonstrated worse overall memory performance. By contrast, the present study demonstrated a crossover interaction, with superior memory for own-gender faces in both participant groups, in the absence of overall group differences (for similar results, see Wright & Sladden, 2003). We thus may conclude that the present OGB was unaffected by general stimulus or participant group effects, and therefore represents a true “in-group” bias.

Both female and male participants reported similar amounts of contact (in h/week and number of contact persons/week). Moreover, although female participants reported more intense contact towards own-gender persons, this difference did not appear to affect memory performance, as the memory bias towards faces of one’s own gender did not correlate with contact quality towards female versus male persons. Thus, the own-gender bias may largely rely on perceived in- versus out-group membership.

In addition to this pattern, which held for recognition memory in general, an interesting gender difference was observed with respect to ‘remember’ responses specifically: the OGB in male but not female participants was accompanied by more remember responses, and thus increased recollection, for own- as compared to other-gender faces. Similarly, previous studies on the own-race bias reported increased recollection for own- relative to other-race faces (Herzmann et al., 2011; Meissner, Brigham, & Butz, 2005). The present findings suggest that increased recollection may not exclusively be related to increased perceptual expertise (see Herzmann & Curran, 2011), but may similarly occur for social categorization biases. Interestingly, male but not female participants in the present study also

rated male faces as more distinctive than female faces. We thus suggest that distinctiveness enhanced recollection of study phase detail.

Event-related potentials during learning

During learning, male participants exhibited a larger N170 for female vs. male faces over the left hemisphere, whereas no corresponding difference was seen in female participants. This finding is somewhat discrepant to previous studies that did not detect effects of face gender on N170 (Mouchetant-Rostaing & Giard, 2003; Mouchetant-Rostaing et al., 2000 but see Ito & Urland, 2003 for larger VPP amplitudes for male relative to female faces). However, previous studies observed a small effect of adaptation to facial gender on the left N170, with smaller amplitudes after adaptation (Kloth, Schweinberger, & Kovács, 2010; see also Kovács et al., 2006). A critical difference between the present study and previous work may concern task requirements: While in previous studies participants were asked to exclusively process face gender, we additionally instructed our participants to memorize the facial stimuli, which may have led to stronger motivation to individuate the faces. In line with this interpretation, a very recent study observed larger N170 amplitudes for racial out-group faces only when facial identity was task-relevant (Senholzi & Ito, in press.). Accordingly, it appears plausible that the effect of face gender on N170 would have not occurred in a passive viewing or simple categorization task. Of note, while previous studies did not investigate interactions between participant and face gender, we detected face gender effects in male participants only. Interestingly, an fMRI study revealed increased activation of the left-hemispheric amygdala and anterior temporal cortex regions for female relative to male faces in male participants, with no corresponding effects in female participants (Fischer et al., 2004).

The finding of a left-hemispheric mechanism discriminating “in-group” from “out-group” faces in the N170 time range is reminiscent of a similar effect in a recent study on the

own-race bias from our group (Wiese et al., in press). In this study, larger N170 amplitudes for other-race faces were observed, and the magnitude of this effect correlated significantly with the behavioural memory bias. In contrast to these results, the present N170 effect for male participants did not correlate significantly with the memory bias ($r = .05, p = .82$), which is unlikely to result from a lack of statistical power, given the complete absence of an effect despite the fact that the number of participants and trials was similar across studies. We thus conclude that there is hardly any evidence that the own-gender bias is moderated by early perceptual processes as reflected in the N170. Effects of subsequent memory were detected in the occipitotemporal P2 (and polarity-reversed in the central N200). P2 for subsequently remembered vs. forgotten male faces was more positive in male participants, while female participants demonstrated more positive amplitudes for subsequently forgotten relative to remembered male faces. The P2 has been interpreted to reflect perceived typicality of faces (Kaufmann & Schweinberger, 2012; Stahl et al., 2008). Accordingly, P2 has been found to be less positive for caricatures, which are more distinctive than veridicals, and similarly for naturally distinctive versus non-distinctive faces (Schulz, Kaufmann, Kurt, & Schweinberger, 2012). According to the multidimensional face space model (Valentine, 1991) distinctive faces are less densely clustered and therefore harder to categorize on dimensions such as race, age, or gender than typical faces. In the present study, female participants may have been more accurate at subsequently remembering particularly distinctive male faces, as a categorization of these faces as belonging to a social out-group might have been delayed, which could have resulted in individuation to some extent. Male participants, by contrast, may have been more accurate at subsequently remembering less distinctive or more typical male faces, perhaps because these are easier to categorize as in-group faces, which according to socio-cognitive models should result in more in-depth processing. Although this interpretation is somewhat tentative at present, it is generally in line with the observed negative correlation

between P2 amplitude and distinctiveness ratings, demonstrating an overall relationship between the two variables.

Finally, a prominent long-lasting Dm effect was observed, but this Dm did not differentiate between subsequently remembered own- and other-gender faces.

Event-related potentials at test

In the test phase of the present study, a larger P2 was detected for female as compared to male faces. While no previous studies examined effects of face gender on P2, Ito and Urland (Ito & Urland, 2003) observed a larger parietal N200 for female faces in a similar time range using a mastoid reference. This finding may be seen in line with the present results, as with an average reference (as used in the present study) a parietal N200 effect may be observed with reversed polarity at occipito-temporal channels. While the processes underlying this effect are not entirely clear at present, it is presumably not related to the OGB in memory, as it similarly occurred in both participant groups.

The first ERP component yielding significant differences between correctly remembered and correctly rejected new items was the N250, which was more negative for hits vs. correct rejections. This finding is generally consistent with previous studies that demonstrated effects of face learning on the N250 (Kaufmann et al., 2009; Tanaka et al., 2006). Moreover, both own-age and own-race bias studies demonstrated additional effects in the N250 range, with more negative amplitudes for other-age and other-race relative to own-group faces, independent of whether faces were learned or new items (Stahl et al., 2010; Wiese, 2012). Such effects may be related to the Early Posterior Negativity (EPN) and may represent the capture of attentional resources by perceptually salient stimuli (Schupp et al., 2007). The absence of a corresponding face gender effect in the N250 time range may reflect the substantial frequency with which other-gender faces are observed.

Finally, larger old/new effects were detected for own-gender faces in female participants. The old/new effect has been interpreted as reflecting conscious recollection, and is known to increase with the amount of study phase detail available (Vilberg & Rugg, 2007). Thus, the present results suggest that female participants not only remember *more* female as compared to male faces, but also recollect more study phase *detail* for own-gender faces. Our finding of an own-gender bias in the amount of ‘remember’-responses in male but not female participants, however, challenges this interpretation. In this context, it should be noted that previous studies on face memory did not observe the typical distribution of old/new effects known from studies on memory for words (see e.g., Rugg & Curran, 2007), with frontal effects representing familiarity and (left) parietal effects reflecting recollection-based recognition. Examining recollection and familiarity in face memory, MacKenzie and Donaldson (2007) found a posterior effect of familiarity (300-500 ms), while a later (500-700 ms) and more anterior effect was related to recollection. The time window used in the present study to analyze the old/new effect overlapped with both segments reported in this previous study, and could thus represent a combination of both processes. Consequently, an analysis on “remember”-responses versus CR was calculated, which more unambiguously reflects recollection-based recognition. This analysis, however, revealed no own-gender bias in the ERP old/new effect. The present ERP findings thus do not provide evidence for enhanced recollection-based recognition of own-gender faces.

Conclusions

Here we observed superior memory for own-gender faces in both male and female participants. This behavioural effect was accompanied by differences due to subsequent memory for male faces in the P2/N200 during learning, which could reflect more efficient categorization of typical male faces as out-group in female, but less efficient categorization as in-group by male participants. Since the own-gender bias was found to be unrelated to

measures of contact, we suggest that these ERP effects do not reflect differences in perceptual expertise, but rather reflect neural correlates of the influence of perceived social category membership on face recognition memory. Overall, our findings suggest that the own-gender bias is mainly related to encoding processes, which is in line with socio-cognitive accounts.

Funding

This work was supported by a grant of the Deutsche Forschungsgemeinschaft (DFG) to H.W. and S.R.S. (grant number WI 3219/5-2).

Acknowledgments

The authors are grateful for help during data collection by Kathrin Rauscher and for support during the Dm analyses by Franziska Krahmer.

References

- Agneta, H., & Johanna, L. (2009). Sex Differences in Cognitive Functions. *Acta Psychologica Sinica*, 41(11), 1081–1090. doi:10.3724/SP.J.1041.2009.01081
- Bartlett, J. C., & Fulton, A. (1991). Familiarity and recognition of faces in old age. *Memory & Cognition*, 19(3), 229–238.
- Bartlett, J. C., & Leslie, J. E. (1986). Aging and memory for faces versus single views of faces. *Memory & Cognition*, 14(5), 371–381.
- Bentin, S., Allison, T., Puce, A., Perek, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience*, 8(6), 551–556.
- Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science*, 18(8), 706–712. doi:10.1111/j.1467-9280.2007.01964.x
- Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods*, 42(1), 286–291. doi:10.3758/BRM.42.1.286
- Byatt, G., & Rhodes, G. (2004). Identification of own-race and other-race faces: Implications for the representation of race in face space. *Psychonomic Bulletin & Review*, 11(4), 735–741.
- Bötzel, K., Schulze, S., & Stodieck, S. R. G. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. *Experimental Brain Research*, 104, 135–143.

- Caharel, S., Montalan, B., Fromager, E., Bernard, C., Lalonde, R., & Mohamed, R. (2011). Other-race and inversion effects during the structural encoding stage of face processing in a race categorization task: An event-related brain potential study. *International Journal of Psychophysiology*, 79(2), 266–271. doi:10.1016/j.ijpsycho.2010.10.018
- Caldara, R., Thut, G., Servois, P., Michel, C. M., Bovet, P., & Renault, B. (2003). Face versus non-face object perception and the “ other-race ” effect: a spatio-temporal event-related potential study. *Clinical Neurophysiology*, 114, 515–528. doi:10.1016/S1388-2457(02)00407-8
- Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). Faces - A database of facial expressions in young , middle-aged, and older women and men: development and validation. *Behavior Research Methods*, 42(1), 351–362.
- Eimer, M. (2011). The face-sensitive N170 component of the event-related brain potential. *The Oxford Handbook of Face Perception*. Oxford University Press (pp. 329–344).
- Fischer, H., Sandblom, C. A. J., Herlitz, A., Fransson, P., Wright, C. I., & Bäckman, L. (2004). Sex-differential brain activation during exposure to female and male faces. *NeuroReport*, 15(2), 235–238. doi:10.1097/01.wnr.0000109795.20952.30
- Furl, N., Phillips, P. J., & O’Toole, A. J. (2002). Face recognition algorithms and the other-race effect: computational mechanisms for a developmental contact hypothesis. *Cognitive Science*, 26(6), 797–815. doi:10.1016/S0364-0213(02)00084-8
- Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and Knowing. *The Oxford Handbook of Memory*. Oxford University Press (pp. 229–244).

- Guillem, F., & Mograss, M. (2005). Gender differences in memory processing: Evidence from event-related potentials to faces. *Brain and Cognition*, 57(1), 84–92.
doi:10.1016/j.bandc.2004.08.026
- Halit, H., De Haan, M., & Johnson, M. H. (2000). Modulation of event-related potentials by prototypical and atypical faces. *NeuroReport*, 11(9), 1871–1875.
- Herzmann, G., & Curran, T. (2011). Experts' memory: An ERP study of perceptual expertise effects on encoding and recognition. *Memory & Cognition*, 39, 412–432.
doi:10.3758/s13421-010-0036-1
- Herzmann, G., Willenbockel, V., Tanaka, J. W., & Curran, T. (2011). The neural correlates of memory encoding and recognition for own-race and other-race faces. *Neuropsychologia*, 49(11), 3103–3115. doi:10.1016/j.neuropsychologia.2011.07.019
- Hugenberg, K., Miller, J., & Claypool, H. M. (2007). Categorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem. *Journal of Experimental Social Psychology*, 43(2), 334–340. doi:10.1016/j.jesp.2006.02.010
- Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The Categorization-Individuation Model: An Integrative Account of the Other-Race Recognition Deficit. *Psychological Review*, 117(4), 1168–1187. doi:10.1037/a0020463
- Ito, T., & Bartholow, B. D. (2009). The neural correlates of race. *Trends in Cognitive Sciences*, 13(12), 524–31. doi:10.1016/j.tics.2009.10.002
- Ito, T., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, 85(4), 616–626. doi:10.1037/0022-3514.85.4.616

- Jeffreys, D. A. (1989). Experimental Brain Research. *Brain Research*, 193–202.
- Kaufmann, J. M., & Schweinberger, S. R. (2012). The faces you remember: Caricaturing shape facilitates brain processes reflecting the acquisition of new face representations. *Biological Psychology*, 89(1), 21–33. doi:10.1016/j.biopsycho.2011.08.011
- Kaufmann, J. M., Schweinberger, S. R., & Burton, A. M. (2009). N250 ERP correlates of the acquisition of face representations across different images. *Journal of Cognitive Neuroscience*, 21(4), 625–641. doi:10.1162/jocn.2009.21080
- Kloth, N., Schweinberger, S. R., & Kovács, G. (2010). Neural correlates of generic versus gender-specific face adaptation. *Journal of Cognitive Neuroscience*, 22(10), 2345–2356. doi:10.1162/jocn.2009.21329
- Kovács, G., Zimmer, M., Bankó, É., Harza, I., Antal, A., & Vidnyánszky, Z. (2006). Electrophysiological correlates of visual adaptation to faces and body parts in humans. *Cerebral Cortex*, 16(5), 742–753. doi:10.1093/cercor/bhj020
- Latinus, M., & Taylor, M. J. (2006). Face processing stages: Impact of difficulty and the separation of effects. *Brain Research*, 1123(1), 179–187. doi:10.1016/j.brainres.2006.09.031
- Lewin, C., & Herlitz, A. (2002). Sex differences in face recognition - Women's faces make the difference. *Brain and Cognition*, 50, 121–128. doi:10.1016/S0278-2626(02)00016-7
- Lovén, J., Herlitz, A., & Rehnman, J. (2011). Women's own-gender bias in face recognition memory: The role of attention at encoding. *Experimental Psychology*, 58(4), 333–340. doi:10.1027/1618-3169/a000100

- Lucas, H. D., Chiao, J. Y., & Paller, K. A. (2011). Why some faces won't be remembered : Brain potentials illuminate successful versus unsuccessful encoding for same-race and other-race faces. *Frontiers in Human Neuroscience*, 5, 1–17.
doi:10.3389/fnhum.2011.00020
- MacKenzie, G., & Donaldson, D. I. (2007). Dissociating recollection from familiarity: electrophysiological evidence that familiarity for faces is associated with a posterior old/new effect. *NeuroImage*, 36(2), 454–463. doi:10.1016/j.neuroimage.2006.12.005
- McKelvie, S. J., Standing, L., St Jean, D., & Law, J. (1993). Gender differences in recognition memory for faces and cars: Evidence for the interest hypothesis. *Bulletin of the Psychonomic Society*, 31(5), 447–448.
- Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7(1), 3–35. doi:10.1037//1076-8971.7.1.3
- Meissner, C. A., Brigham, J. C., & Butz, D. A. (2005). Memory for own- and other-race faces: A dual-process approach. *Applied Cognitive Psychology*, 19(5), 545–567.
doi:10.1002/acp.1097
- Mercure, E., Dick, F., & Johnson, M. H. (2008). Featural and configural face processing differentially modulate ERP components. *Brain Research*, 1239, 162–170.
doi:10.1016/j.brainres.2008.07.098
- Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, 36(4), 630–633.

Mouchetant-Rostaing, Y., & Giard, M. H. (2003). Electrophysiological correlates of age and gender perception on human faces. *Journal of Cognitive Neuroscience*, *15*(6), 900–910.

doi:10.1162/089892903322370816

Mouchetant-Rostaing, Yolande, Giard, M.-H., Bentin, S., Aguera, P.-E., & Pernier, J. (2000).

Neurophysiological correlates of face gender processing in humans. *European Journal of Neuroscience*, *12*, 303–310.

Paller, K A, Kutas, M., & Mayes, A. R. (1987). Neural correlates of encoding in an incidental learning paradigm. *Electroencephalography and Clinical Neurophysiology*, *67*(4), 360–371.

Paller, K.A., McCarthy, G., & Wood, C. C. (1988). ERPs predictive of subsequent recall and recognition performance. *Biological Psychology*, *26*, 269–276.

Phillips, P. J., Wechsler, H., Huang, J., & Rauss, P. J. (1998). The Feret database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, *16*, 295–306.

Rehnman, J., & Herlitz, A. (2006). Higher face recognition ability in girls: Magnified by own-sex and own-ethnicity bias. *Memory*, *14*(3), 289–296. doi:10.1080/09658210500233581

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin*, *138*(1), 146–174.

doi:10.1037/a0025750

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. doi:10.1016/j.tics.2007.04.004

- Schulz, C., Kaufmann, J. M., Kurt, A., & Schweinberger, S. R. (2012). Faces forming traces: neurophysiological correlates of learning naturally distinctive and caricatured faces. *NeuroImage*, 63(1), 491–500. doi:10.1016/j.neuroimage.2012.06.080
- Schupp, H. T., Stockburger, J., Bublatzky, F., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2007). Explicit attention interferes with selective emotion processing in human extrastriate cortex. *BMC Neuroscience*, 8(16), 1–12. doi:10.1186/1471-2202-8-16
- Schweinberger, S. R., & Burton, A. M. (2003). Covert Recognition and the neural system for face processing. *Cortex*, 9–30.
- Schweinberger, S. R., Pfütze, E.-M., & Sommer, W. (1995). Repetition Priming and Associative Priming of Face Recognition: Evidence from Event-Related Potentials. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21(3), 722–736.
- Schweinberger, S. R., Pickering, E. C., Jentsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research*, 14(3), 398–409.
- Senholzi, K. B., & Ito, T. (in press). Structural face encoding: How task affects the N170's sensitivity to race. *Social Cognitive and Affective Neuroscience*, 1–6.
doi:10.1093/scan/nss091
- Sommer, W., Heinz, A., Leuthold, H., Matt, J., & Schweinberger, S. R. (1995). Metamemory, distinctiveness, and event-related potentials in recognition memory for faces. *Memory & Cognition*, 23(1), 1–11.

- Sommer, W., Schweinberger, S. R., & Matt, J. (1991). Human brain potential correlates of face encoding into memory. *Electroencephalography and Clinical Neurophysiology*, 79(6), 457–463.
- Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, 7(1), 36–97. doi:10.1037//1076-8971.7.1.36
- Stahl, J., Wiese, H., & Schweinberger, S. R. (2008). Expertise and own-race bias in face processing: an event-related potential study. *NeuroReport*, 19(5), 583–587.
doi:10.1097/WNR.0b013e3282f97b4d
- Stahl, J., Wiese, H., & Schweinberger, S. R. (2010). Learning task affects ERP-correlates of the own-race bias, but not recognition memory performance. *Neuropsychologia*, 48(7), 2027–2040. doi:10.1016/j.neuropsychologia.2010.03.024
- Sun, Y., Gao, X., & Han, S. (2010). Sex differences in face gender recognition : An event-related potential study. *Brain Research*, 1327, 69–76. doi:10.1016/j.brainres.2010.02.013
- Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of preexisting and acquired face representations: The N250 event-related potential as an index of face familiarity. *Journal of Cognitive Neuroscience*, 18(9), 1488–1497.
doi:10.1162/jocn.2006.18.9.1488
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology*, 2(43A), 161–204.
- Valentine, T. (1992). Towards an exemplar model of face processing: the effects of race and distinctiveness. *The Quarterly Journal of Experimental Psychology*, 4(44A), 671–703.

- Vilberg, K. L., & Rugg, M. D. (2007). Dissociation of the neural correlates of recognition memory according to familiarity, recollection, and amount of recollected information. *Neuropsychologia*, 45(10), 2216–25. doi:10.1016/j.neuropsychologia.2007.02.027
- Walker, P. M., Silvert, L., Hewstone, M., & Nobre, A. C. (2008). Social contact and other-race face processing in the human brain. *Social Cognitive and Affective Neuroscience*, 3(1), 16–25. doi:10.1093/scan/nsm035
- Wiese, H. (2012). The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. *Biological Psychology*, 89(1), 137–147. doi:10.1016/j.biopsycho.2011.10.002
- Wiese, H., Kaufmann, J. M., & Schweinberger, S. R. (in press). The Neural Signature of the Own-Race Bias: Evidence from Event-Related Potentials. *Cerebral Cortex*. doi:10.1093/cercor/bhs369
- Wiese, H., Komes, J., & Schweinberger, S. R. (2012). Daily-life contact affects the own-age bias and neural correlates of face memory in elderly participants. *Neuropsychologia*, 50(14), 3496–3508. doi:10.1016/j.neuropsychologia.2012.09.022
- Wiese, H., Schweinberger, S. R., & Hansen, K. (2008). The age of the beholder: ERP evidence of an own-age bias in face memory. *Neuropsychologia*, 46(12), 2973–2985. doi:10.1016/j.neuropsychologia.2008.06.007
- Wiese, H., Stahl, J., & Schweinberger, S. R. (2009). Configural processing of other-race faces is delayed but not decreased. *Journal of Experimental Psychology*, 81, 103–109. doi:10.1016/j.biopsycho.2009.03.002

- Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica*, 114, 101–114. doi:10.1016/S0001-6918(03)00052-0
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46(3), 441–517.
doi:10.1006/jmla.2002.2864
- Young, S. G., Bernstein, M. J., & Hugenberg, K. (2010). When Do Own-Group Biases in Face Recognition Occur? Encoding versus Post-Encoding. *Social Cognition*, 28(2), 240–250.

Table 1

Performance measures from study and test phases

| | Female Participants | | Male Participants | |
|-------------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | Female Faces | Male Faces | Female Faces | Male Faces |
| Study Phases | | | | |
| RT \pm SD | 843.11 \pm 239.89 | 832.62 \pm 210.97 | 888.79 \pm 222.57 | 869.41 \pm 208.51 |
| ACC \pm SD | 0.99 \pm 0.02 | 0.99 \pm 0.02 | 0.98 \pm 0.02 | 0.98 \pm 0.02 |
| Test Phases | | | | |
| RT \pm SD (hits) | 1053.48 \pm 143.24 | 1076.38 \pm 154.64 | 1053.35 \pm 118.60 | 1040.92 \pm 101.26 |
| RT \pm SD (false alarms) | 854.64 \pm 270.65 | 874.40 \pm 348.18 | 902.91 \pm 374.42 | 883.75 \pm 344.28 |
| hits \pm SD | 0.74 \pm 0.14 | 0.65 \pm 0.16 | 0.62 \pm 0.12 | 0.66 \pm 0.13 |
| false alarms \pm SD | 0.20 \pm 0.14 | 0.20 \pm 0.15 | 0.17 \pm 0.11 | 0.17 \pm 0.12 |
| remember (hits) | 0.47 \pm 0.22 | 0.47 \pm 0.26 | 0.43 \pm 0.20 | 0.49 \pm 0.20 |
| know (hits) | 0.21 \pm 0.15 | 0.18 \pm 0.15 | 0.23 \pm 0.14 | 0.23 \pm 0.15 |
| guess (hits) | 0.02 \pm 0.03 | 0.02 \pm 0.03 | 0.02 \pm 0.03 | 0.03 \pm 0.03 |
| d' \pm SD | 1.69 \pm 0.73 | 1.48 \pm 0.67 | 1.43 \pm 0.54 | 1.60 \pm 0.61 |
| C \pm SD | 0.13 \pm 0.32 | 0.31 \pm 0.42 | 0.32 \pm 0.30 | 0.27 \pm 0.39 |

Note: means \pm standard deviation

Table 2

Face Rating: attractiveness and distinctiveness

| Attractiveness | Participants | |
|-----------------|-----------------------------|-----------------------------|
| | <u>Female</u> | <u>Male</u> |
| Faces | | |
| Female | 2.83 ± 0.59 | 2.51 ± 0.55 |
| Male | 2.56 ± 0.57 | 2.42 ± 0.52 |
| | Mean attractiveness rating | Mean attractiveness rating |
| | 2.69 ± 0.56 | 2.46 ± 0.50 |
| Distinctiveness | Participants | |
| | <u>Female</u> | <u>Male</u> |
| Faces | | |
| Female | 3.20 ± 0.56 | 3.19 ± 0.78 |
| Male | 3.22 ± 0.52 | 3.45 ± .078 |
| | Mean distinctiveness rating | Mean distinctiveness rating |
| | 3.21 ± 0.52 | 3.32 ± 0.76 |

Note: means ± standard deviation

Table 3

Contact Data

| Variable | Participants | Contact persons | |
|-------------------|--------------|-----------------|---------------|
| <i>Daily Life</i> | Females | Females | Males |
| Contact time | | 50.20 ± 38.16 | 49.58 ± 26.38 |
| Contact numbers | | 14.46 ± 5.68 | 13.96 ± 7.54 |
| Contact quality | | 3.03 ± 0.62 | 2.03 ± 0.46 |
| Variable | Participants | Contact persons | |
| <i>Daily Life</i> | Males | Females | Males |
| Contact time | | 62.04 ± 68.27 | 67.86 ± 94.54 |
| Contact numbers | | 22.21 ± 15.26 | 26.91 ± 16.65 |
| Contact quality | | 2.30 ± 0.56 | 2.27 ± 0.67 |

Note: means ± standard deviation

Figure Captions

Figure 1. Sensitivity (d') values depicting recognition performance in the test phases, separate for female and male participants and female and male face gender. Error bars depict standard errors of the mean.

Figure 2. Upper Part: Grand Mean ERPS during learning for female and male participants and female and male faces separately. Vertical lines demarcate occipito-temporal P2 (220 - 270) and central N200 (220-270), the N250 (from 270-400 ms) and late Dm (from 450 – 1000 ms) time windows. Lower Part: Voltage maps of ERP differences (sub. remembered minus sub. forgotten) in the P2/N200 time window are depicted (spherical spline interpolation, 90° equidistant projection).

Figure 3. Upper Part: Grand mean ERP old/new effects (hits vs. CR) separately for male and female participants and female and male faces. Vertical lines at 400 and 675 ms demarcate the old/new effect as evaluated statistically. Additionally vertical lines at P9/P10 demarcate P2 (220 - 270) and the N250 (from 270-400 ms) time windows during test. In the lower part voltage maps of ERP differences (hits minus CR) in the old/new effect time window are depicted (spherical spline interpolation, 90° equidistant projection).

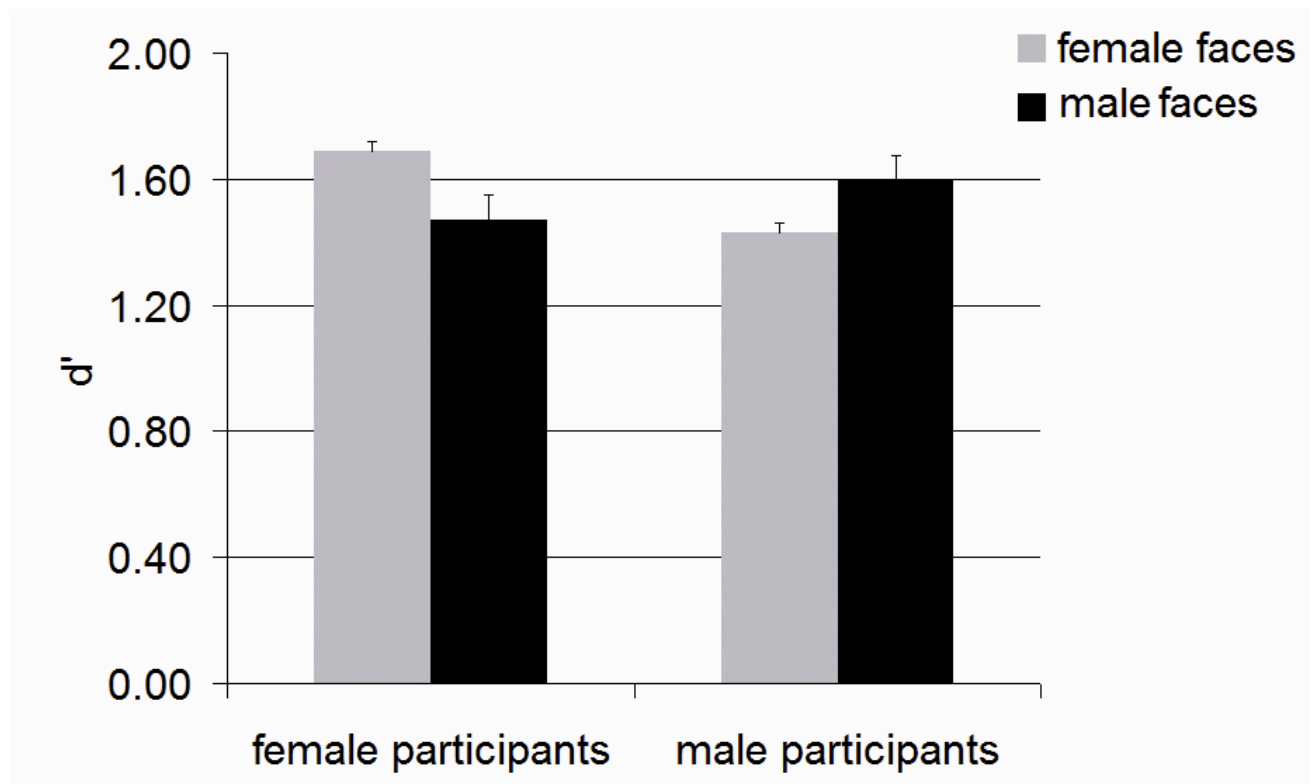
Figure 1

Figure 2

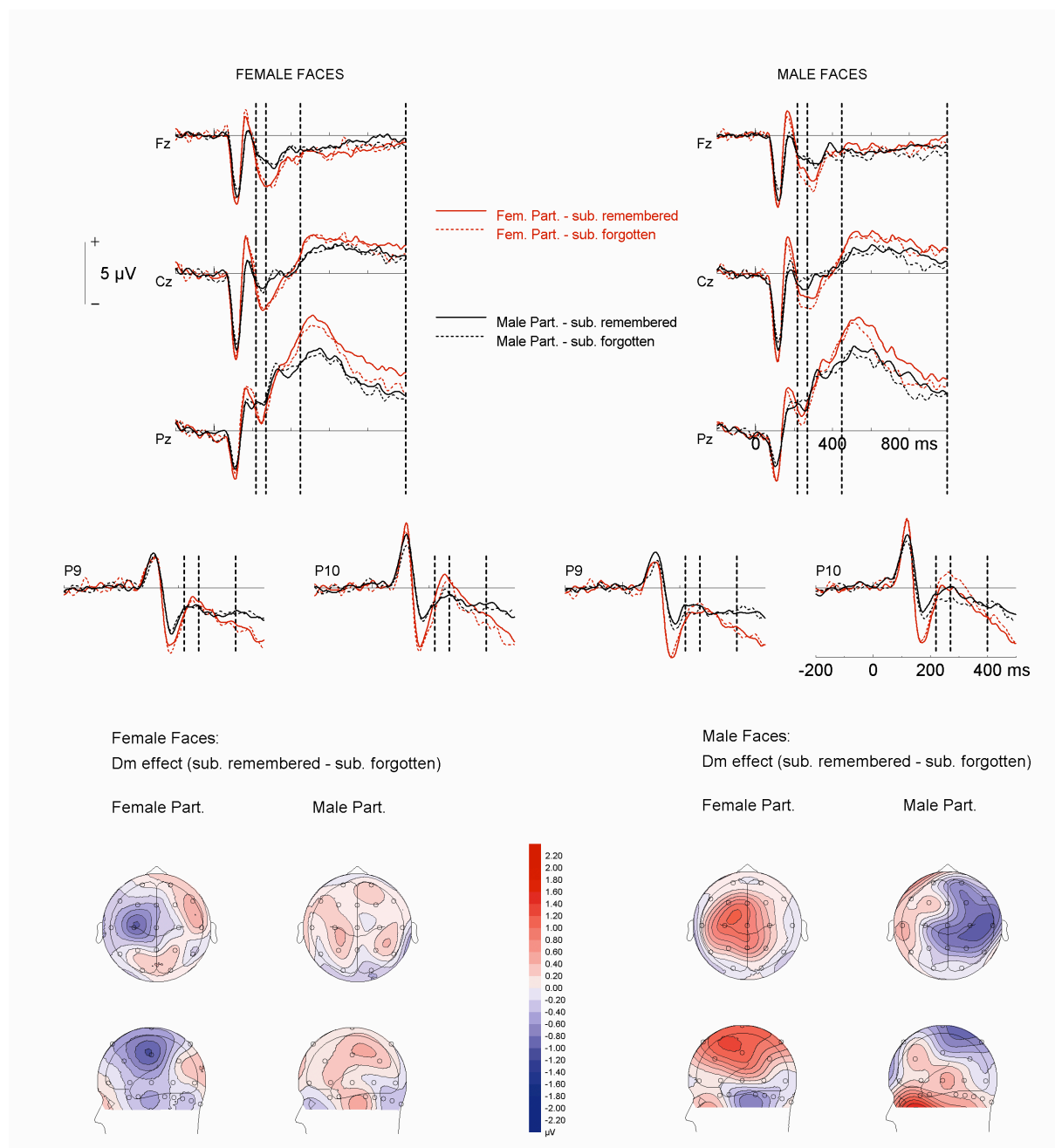


Figure 3

